Illuminating the Instructional Experience

Centering Students in Education Measurement

Leading Educators

Potential, ignited.

Illuminating the Instructional Experience
Centering Students in Education Measurement

June 2021

Acknowledgements

Introduction

Over the past two decades, there has been an increasing focus on measuring outcomes in education. Families care deeply about their child's education. Policymakers want to understand whether investments in education are working. Community members know that strong local education options matter to collective wellbeing and economic health. Thus, schools and school systems have incentives to gather and report data, adopt evidence-based interventions, and demonstrate impact. Unfortunately, many of the most common evaluation methods are incomplete at best, limiting our ability to make smart changes to teaching and learning.

Many accountability metrics create a culture of what Bettina Love calls education survival, which holds students accountable for systemic injustices while missing what helps them thrive: humanity, dignity, and the ability to fight for justice in their communities.[1] These accountability metrics are also limited in other important ways. For example, observation rubrics and studies are often content-agnostic,[2] poorly aligned to state-standards,[3] or narrowly focused on one dimension of instruction.[4] In order to reach educational equity, we must re-examine measurement tools through a critical, anti-racist lens. Furthermore, this measurement work must be coupled with active efforts to disrupt and transform the systems, policies, and practices that perpetuate racism in our schools and communities.

Beginning in 2009 with the Race to the Top initiative, accountability pressure intensified as states increasingly moved to adopt rigorous nationally-aligned standards to prepare students for a rapidly changing workforce.[5] But while many teachers today believe they are teaching to these standards, research shows many teachers do not yet deeply understand the key shifts in mathematics and English Language Arts.[6] Without this knowledge, teachers may unintentionally create lessons that lack rigor and miss key concepts students need for college and career. For example, Education Trust examined classroom tasks and found that only 38% of classroom tasks met grade-level expectations,[7] and TNTP found that students spent the equivalent of six months per school year on assignments that were below grade-level.[8]

These challenges can be exacerbated by teachers' implicit biases and the varying expectations that they have for students of different backgrounds and identities. Research has demonstrated that teachers tend to have higher expectations for students who share their own racial identities[9] and that those expectations matter for students' achievement both within school and beyond.[10] Lowered

1 Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom*. Beacon Press.

2 Aligning Content and Practice: The Design of the Instructional Practice Guide, Student Achievement Partners (2019).

3 Welch, M.J., Davis, S., Isaia, R.G., Johnston, W.T., Stein, L.B., Jenuwine, H., & Macdonald, K. (2016). Aligning evaluation: How much do teacher evaluation rubrics emphasize Common Core?

4 Ball, D. L., & Forzani, F. M. (2007). What Makes Education Research" Educational"?. *Educational Researcher, 36*(9), 529.

5 Boser, U. (2012). Race to the Top: What Have We Learned from the States So Far? A State-by-State Evaluation of Race to the Top Performance. *Center for American Progress*.

6 Kaufman, Julia H., V. Darleen Opfer, Lindsey E. Thompson, and Joseph D. Pane (2018). Connecting What Teachers Know About State English Language Arts Standards for Reading and What They Do in Their Classrooms: Findings from the American Teacher Panel. Santa Monica, CA: RAND Corporation.
Opfer, V. Darleen, Julia H. Kaufman, Joseph D. Pane, and Lindsey E. Thompson, Aligned Curricula and Implementation of Common Core State Mathematics Standards: Findings from the American Teacher Panel. Santa Monica, CA: RAND Corporation, 2018.

7 Ed Trust (2016) "Checking In: Do Classroom Assignments Reflect Today's Higher Standards?" https://edtrust.org/resource/classroomassignments/

8 TNTP (2018). *The Opportunity Myth*: *The Opportunity Myth*: *What Students Can Show Us About How School Is Letting Them Down—and How to Fix It.*. New York, NY: TNTP.

9 Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of education review, 52*, 209-224.

10 Papageorge, N. W., Gershenson, S., & Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics, 102*(2), 234-251.

expectations can lead a teacher to present materials that are below grade level or that do not possess the rigor necessary for college and career readiness. Nearly 80% of the teacher workforce in U.S. public schools identifies as White,[11] whereas more than 53% of the students in K–12 public schools are students of color.[12] This presents a serious challenge to ensuring high expectations for all students, especially if adequate support for recognizing and remedying bias is not in place.

The underrepresentation of people of color in teaching and leadership roles also shapes students' sense of belonging. Decades of school climate research has demonstrated that students who possess a strong sense of belonging in their classrooms and schools fare better on a host of academic, social, and behavioral outcomes.[13]

Unfortunately, debiasing efforts have shown limited effectiveness,[14] and while diversifying the teaching workforce is an urgent priority, improving the recruitment and retention of teachers of color is a complex, multi-faceted challenge requiring change in structural, institutional, programmatic, and environmental factors.[15] In addition to the necessary but long-term work for systemic change, near-term strategies to advance the quality of teaching can dramatically improve students' experience and academic outcomes.[16]

Culturally responsive pedagogies have been put forth as one means of addressing these different opportunity, expectations, and representation gaps. Geneva Gay defines culturally responsive pedagogy using the following criteria: "it teaches to and through

[students'] strengths . . . it builds bridges of meaningfulness between home and school experiences as well as between academic abstractions and lived sociocultural realities; it uses a wide variety of instructional strategies that are connected to different learning styles; it incorporates multicultural information, resources, and materials in all the subjects and skills routinely taught in schools."[17] Gloria Ladson-Billings's work with culturally responsive pedagogies emphasizes the importance of critical consciousness, or of not simply incorporating cultural examples into instruction but also pushing students to critically engage with issues that directly impact their lives and communities.[18] Both of these frameworks acknowledge that a primary aim of culturally responsive pedagogies is to recognize and honor the diverse cultural heritages and experiences that students bring with them into the classroom and to use those as a resource for teaching and learning. This is particularly important for collective well-being in a pluralistic, multicultural society like the United States.

To improve outcomes for all students, we believe teaching must integrate pedagogical content knowledge, cultural responsiveness, and well-being, but these themes are often addressed in isolation. Social emotional well-being and academics are often taught divorced from the sociopolitical context of injustice, inequity, and violence that makes up the day-to-day lives of students.[19] At the same time, each of these issues is incredibly complex, with each domain being the subject of decades of scholarship and intervention efforts, so school systems may find it difficult or unfeasible to address all three at once.

11 National Center for Education Statistics (2020). *Race and ethnicity of public school teachers and their students* (NCES 2020-103). Washington, DC: U.S. Department of Education.

12 National Center for Education Statistics (2020). *Digest of Education Statistics, Chapter 2* (Table 203.50). Washington, DC: U.S. Department of Education.

13 Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M., & de Boer, H. (2020). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Research Papers in Education, 35*(6), 641-680.

14 Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., … & Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General, 145*(8), 1001.

15 Carter Andrews, D. J., Castro, E., Cho, C. L., Petchauer, E., Richmond, G., & Floden, R. (2019). Changing the narrative on diversifying the teaching workforce: A look at historical and contemporary factors that inform recruitment and retention of teachers of color.

16 Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review, 104*(9), 2633-79.

17 Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of teacher education, 53*(2), 106-116.

18 Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: aka the remix. Harvard Educational Review, 84(1), 74-84.

19 Simmons, D. (2019). Why we can't afford whitewashed social-emotional learning. *ASCD Education Update, 61*(4), 2-3.

A common strategy is to take a layered approach to build knowledge and capacity over a set number of years, but narrowing focus on one aspect at a time runs the risk of reinforcing the systemic inequities leaders seek to address.[20] For example, strategies to support social emotional well-being that do not examine power and privilege or standards-aligned curricula that do not consider cultural representation may reinforce an exclusionary dominant culture. Instead, narrowing must occur within a framework that acknowledges the interdependence of these domains.

To support holistic student success, school systems need instructional support and evaluation strategies that focus on an integrated academic vision while also prioritizing a narrow set of high-leverage targets. While many schools focus on student data and teacher practice analysis, few collect data on culturally responsive practices or practices to support student well-being. There is a pressing need to develop coherent data systems that allow leaders to understand how efforts to improve instruction influence instructional practices, social emotional well-being, student sense of belonging, and student learning. Support partners, like Leading Educators, can play a powerful role in developing such systems to empower leaders to make better-informed decisions.

Deepening our understanding of teacher practice and development

In 2019, Leading Educators recognized the need for help with a critical growth area. Four years after making a shift to content- and curriculum-specific professional learning, we had strong systems for coaching and development as well as useful tools and practices for measuring changes in knowledge, beliefs, conditions, and student outcomes. We had supported thousands of educators to strengthen teaching in ongoing cycles of learning, application, and data analysis.

Furthermore, a rigorous quasi-experimental study indicated our approach led to significant positive effects on both ELA and math student achievement.[21] When it came to measuring change in instructional practice, however, we lacked the evidence that our partners needed for continuous improvement.

We used formative systems for observation and teacher feedback focused on the standards and content, but we did not always visit the same classrooms, use the same tool or section of a tool, or devote time to strong calibration. As a result, the evidence we had was best suited as feedback for individual teachers or schools but was not yet reliable or able to assess change over time. Additionally, we had not yet integrated social and emotional well-being and culturally responsive pedagogies as focal points for observations.

After consulting with research literature, peer organizations, and funders, we recognized that this was a sector-wide problem. For instance, in a 2016 review by the American Institute of Research of 45 instructional observation rubrics from nonprofits, local education agencies, and state education agencies found that 82 percent of rubrics were content- and grade-band agnostic and were generally poorly aligned to college and career readiness standards.[22]

Like many of our peer organizations, we were using the Instructional Practice Guide designed by Student Achievement Partners, but the tool is not intended for research or evaluation and does not integrate culturally responsive teaching practices.[23] Additionally, while practitioners and scholars have used culturally responsive pedagogies for decades, the effectiveness of different approaches continues to be an nascent area of inquiry within educational research. A small but growing body of empirical research has begun to generate meaningful knowledge and evidence of the positive impact of culturally responsive pedagogies using qualitative methods. The education sector must continue to develop and align on strong measurement tools, especially

20 Leading Educators (2021). *Teaching for Equity. https://leadingeducators.org/equity/*
21 Leading Educators (2019). Understanding Leading Educators Impact on Louisiana and Michigan Students. Available from https://leadingeducators.org/news-resource/brightspots/
22 Welch, M.J., Davis, S., Isaia, R.G., Johnston, W.T., Stein, L.B., Jenuwine, H., & Macdonald, K. (2016). Aligning evaluation: How much do teacher evaluation rubrics emphasize Common Core
23 Aligning Content and Practice: The Design of the Instructional Practice Guide, Student Achievement Partners (2019).

at the teacher level, to build a robust evidence base.[24]

To address this data gap, we convened research and practice experts to engage in a year and a half of collaboration. Ultimately, we hoped to improve our own practices to strengthen evidence collection and uncover generalizable principles to share across the education sector. To achieve those goals, we shared our data, our calibration processes, and our challenges with a board of diverse experts. This report shares our story of collaboration, the lessons learned along the way, and the focus questions we pose to the sector as we work toward educational equity. We hope these will be useful to anyone who conducts classroom observations or seeks to achieve equitable outcomes for students.

Building an advisory team

In February 2020, we began to recruit five members for an Evaluation Advisory Board to review our current data practices and provide recommendations for improving methods, implementation, and analysis. We looked for members with strong backgrounds in educational research who would also bring diverse methodological approaches (e.g., measurement, evaluation, qualitative data analysis, and economics) and perspectives. They included experts in both ELA and math content, current and former educators, and academics and non-profit practitioners.

- **Dr. Robert Q. Berry, III,** *President of the National Council of Teachers of Mathematics and Professor in the Curry School of Education at University of Virginia*: A former mathematics teacher, Dr. Berry teaches elementary and special education mathematics methods courses in the teacher education program at the University of Virginia. Additionally, he teaches graduate level mathematics education courses and courses for in-service teachers seeking a mathematics specialist endorsement. His research focuses on equity issues in mathematics education, pre- and in-service teachers'

mathematical knowledge for teaching, and mathematics instructional quality. Berry has extensive experience in classroom observation and has collaborated with other researchers to develop an observation instrument, Mathematics Scan, to examine mathematics teaching quality.

- **Dr. Miah Daughtery,** *Director of Content Advocacy and Design, ELA at NWEA*: A focused and passionate educator, Dr. Miah E. Daughtery has more than fifteen years' experience at the classroom, district, state and national levels. After working as a teacher and district leader in Detroit and Las Vegas, Dr. Daughtery transitioned to the Tennessee Department of Education as the state coordinator for literacy, where she developed systems to ensure content accuracy and alignment for the literacy portion of the state assessment and led statewide literacy professional learning efforts. She joined Achieve in 2016 as the Director, ELA & Literacy, providing technical and advisory support to states, reviewing college-and-career standards revisions for ELA and evaluating state summative assessments. In 2020, she joined NWEA to lead the organization's vision and direction for literacy.

- **Jessica Eadie,** *Director at Student Achievement Partners*: Through her work at Student Achievement Partners, Eadie supports efforts at the classroom, district and state levels to help educators as they enact challenging K-12 academic standards in ELA/literacy and mathematics classrooms. Her work involves bringing learnings from research to practical resources and support for educators with the goal of disrupting inequities and improving learning outcomes for students. She previously served as the Executive Director of Assessment at the New York City Department of Education.

- **Dr. Heather Hill,** *Professor at Harvard University*: Dr. Heather C. Hill studies policies and programs designed to improve mathematics and teaching quality. Her recent research focuses on: teacher

---

24 Bottiani, J. H., Larson, K. E., Debnam, K. J., Bischoff, C. M., & Bradshaw, C. P. (2018). Promoting educators' use of culturally responsive practices: A systematic review of inservice interventions. *Journal of Teacher Education, 69*(4), 367-385.

professional development, instructional coaching, teacher evaluation, changes over time in teachers' mathematical knowledge and instructional quality in mathematics, and the teacher experiences and characteristics that lead to high-quality instruction and stronger student outcomes. Dr. Hill and her team have developed assessments that capture teachers' mathematical knowledge for teaching and teachers' mathematical quality of instruction, assessments now widely available to researchers, instructional coaches, evaluators, and policy-makers via online training and administrative systems. Dr. Hill is a fellow of the American Educational Research Association (AERA), and serves on the AERA grants board, on the editorial boards of several journals, and as an advisor to numerous research projects and policy efforts in both the U.S. and abroad. She is co-author of Learning Policy: When State Education Reform Works with David K. Cohen (Yale Press, 2001).

- **Dr. Matthew Steinberg,** *Associate Professor at George Mason University*: Dr. Steinberg is an Associate Professor of Education Policy in the College of Education and Human Development, a University Affiliate Faculty at the Schar School of Policy and Government, and Director of EdPolicyForward: The Center for Education Policy at George Mason University. Dr. Steinberg is an Affiliated Researcher with the University of Chicago Consortium on School Research, a Faculty Affiliate with the Annenberg Institute for School Reform at Brown University, and an IUR Scholar at the University of Pennsylvania Institute for Urban Research. His research addresses issues of educational significance at the intersection of the economics of education and education policy, including: teacher evaluation and human capital; urban school reform; school discipline and safety; and school finance. His work informs local and national policy discussions on the impact of education policies and practices on the distribution of teacher effectiveness and the educational outcomes of students, particularly the most disadvantaged among the population.

Setting a rhythm for collaboration

After convening the Evaluation Advisory Board, we set a quarterly meeting structure. At each meeting, we presented context-specific challenges as well as problems of practice that cut across all schools and systems we serve. At each meeting, advisors debated, provided feedback, and shared from their own work.

In early 2020 as the COVID-19 pandemic was just beginning, our conversations focused on how to respond to the crisis. How should we rethink our work in light of shifts between virtual and hybrid instruction? What did our partners most need from us to support student engagement and learning? As the year went on, we recognized "normal" would not be returning any time soon, and we needed to adjust to the new realities of schooling for both the short- and long-term.

When we could, we grounded our advisory board conversations in specific evidence. Over the winter and spring, we collected and analyzed a set of instructional artifacts from teaching and learning during the pandemic. Teachers submitted lesson plans with written annotations or videos of instruction along with three to five student work samples from a task in that lesson. We recruited a set of external observers, all current and former school and district leaders with at least three years of experience with the Instructional Practice Guide. After pairs of observers analyzed each artifact set, we brought the data to our advisors and used the subsequent meeting to sharpen our tool and approach, then repeated the process.

Over the course of the year, this work centered around several key questions.

- How could we broaden our observation work to be comprehensive of not just content and standards but also culturally responsive pedagogies and social-emotional well-being?

- How could we move beyond teacher actions to better understand what students were experiencing academically, personally, and culturally?

- How should we focus data collection on the highest leverage constructs for ensuring teachers and students thrive?

Integrating academics, culturally responsive pedagogies, and social and emotional learning

Starting in 2019, we developed *Teaching for Equity*, a framework to help educators see rigorous instruction as critically interwoven with well-being and culturally responsive pedagogies. Only through attending to all three strands can we create schools where all students thrive. As we began to support schools in using the framework, we recognized we needed to create methods for seeing *Teaching for Equity*'s beliefs and practices in action. While researchers have created strong instruments focused on academic content[25] and culturally-responsive observation protocols[26] individually, we did not find a tool that wove them together. Our partners asked for specific language and tools to do just that. At the same time, we recognized the danger of using measurement tools that highly rate standards-aligned instruction while unintentionally perpetuating dominant culture values.

So, we set out to respond to this gap with a new tool. Starting with the beliefs and practices highlighted in *Teaching for Equity*, we identified a set of indicators that assessed the alignment of the content and tasks to grade-level standards and also asked questions about culturally responsive teaching practices and well-being, such as "Do students see/experience relevant historical and current representations of their culture in the instructional content (mirrors)?" Our advisors urged us to ensure that the tool would not simply put these three domains alongside each other without articulating how they fit together. They posed, "What does it mean to value students' cultures and identities in the contexts

of a mathematics or ELA classroom? How can the indicators become more concrete to avoid surface-level references to or even harmful stereotypes of students' cultures?" The advisors urged us to be specific in our articulation of the indicators and to engage in extensive calibration to train raters using both examples and non-examples.

For greater feasibility, we sought to prioritize a small set of practices. We began with 14 indicators that we believe matter most for students, such as practices that may support a sense of belonging and a focus on students' work. For the coming school year, we seek to narrow further to five or six indicators. We knew this choice meant we would not focus on a comprehensive vision of instruction but would instead provide a model for how systems can take narrow and focused steps toward the integrated vision in *Teaching for Equity*. We piloted and revised the tool twice with a modest sample of 19 instructional artifact sets. We recruited a set of external observers with diverse backgrounds, prioritizing both deep experience with standards-based instructional observations and experience with racial equity frameworks.

Overall, we found the artifact sets received higher scores on average on the sections of the tool aligned with college and career readiness but scored lower on indicators connected to cultural relevance and identity. **Figure 1** contrasts two sample indicators to illustrate this trend. As one of our observers reflected after the first round of piloting, "...the questions around content that challenges oppressive narratives, incorporates student funds of knowledge,[27] and values students' culture and identity... are ideas I consider a lot in my own classroom throughout the year, but not as much when giving feedback or when planning a daily lesson." We took this as evidence that the tool was capturing important aspects of instruction (e.g., cultural responsiveness) that

---

25 "Including the subject-specific instruments used in the Measures of Effective Teaching (MET) study, the Mathematical Quality of Instruction and the ELA-focused Protocol for Language Arts Teaching Observation, and the mathematics-focused Teaching for Robust Understanding rubric" - Achieve the Core 2016:2

26 Powell, R., Chambers Cantrell, S., Malo-Juvera, V., & Correll, P. (2016). Operationalizing culturally responsive instruction: Preliminary findings of CRIOP research. *Teachers College Record, 118*(1), 1-46.

27 funds of knowledge - a concept created by Dr. Carlos Vélez-Ibáñez and Dr. James Greenberg to describe the historical accumulation of abilities, bodies of knowledge, assets, and cultural ways of interacting that were evident in U.S.-Mexican households. These funds of knowledge are culturally, socially, and cognitively-complex resources that teachers can use to enhance their students' academic progress in context.

are often overlooked or ignored by traditional, content-focused teacher evaluation measures.

We had two observers review each set of artifacts to test the tool's reliability. Because observers' had experience using the Instructional Practice Guide, we expected stronger rater alignment on indicators focused on standards-based instructional practices relative to indicators focused on culturally responsive pedagogies. To address this, we paid special attention to indicators related to culturally responsive practices in calibration, discussing both examples and nonexamples. Observers rated each indicator as "2 - Strong evidence", "1 - Some evidence" or "0 - No evidence."

When comparing ratings across observer pairs for the same indicator for the same artifact, the expected to see a pattern of lower alignment for indicators related to culturally responsive pedagogies that did not appear. In fact, the average difference between ratings was slightly lower for indicators related to culturally-responsive practices or well-being than it was for indicators related to rigorous academics. In general, the frequency with which raters agreed exactly on ratings was lower than what we had achieved previously with the Instructional Practice Guide (54% versus 62%, respectively), which could be expected with a new tool.

Looking forward, these pilots reaffirmed the importance of balancing all three dimensions of *Teaching for Equity*. Centering observation and feedback or district level planning on this or a similar set of indicators will support coherence across instructional improvement areas as well as integration of key dimensions of student learning and well-being. As observations increasingly focus on a more integrated set of indicators, the need for careful recruitment, selection, and calibration of observers grows even more crucial to ensure alignment and mitigate bias.

Problem	Potential mitigation strategies
Culturally responsive pedagogies, social-emotional well-being, and academics are often seen as separate, prompting educators to focus on one at a time at the expense of coherent, culturally responsive, rigorous instruction. Additionally, social emotional well-being and culturally responsive pedagogies are often considered at a surface level.	• Ensure the observation tool integrates rigorous academics, social-emotional well-being, culturally responsive pedagogies • Select a narrow subset of indicators that are well supported by evidence and tied to balanced instructional priorities • Consider the identities of the observers, the teachers, and those being observed • Be specific and use extensive norming with both examples and non-examples

Figure 1: Comparison of Grade-Level Content and Culturally Responsive Practices Indicators
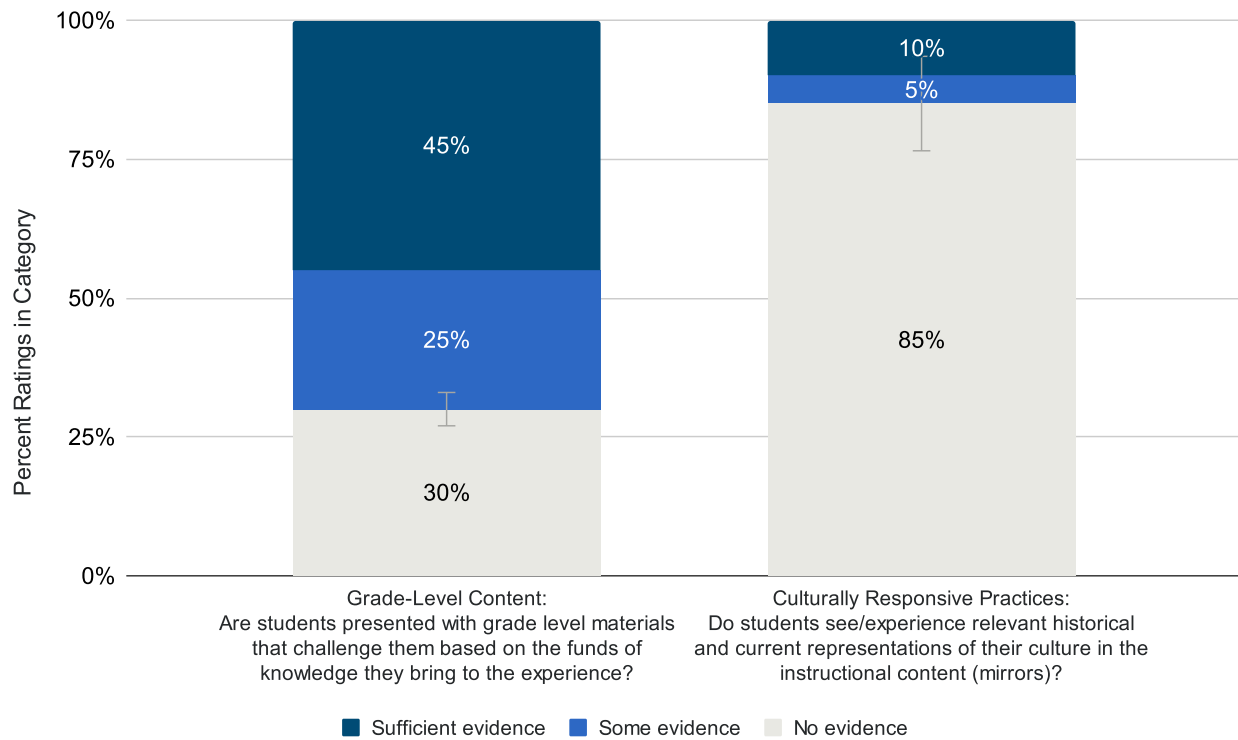
From Leading Educators' second pilot round of instructional artifact set analysis in May 2021. For this round, pairs of external observers independently assessed 10 artifact sets.

Figure 2: Average Absolute Rating Discrepancy by Domain
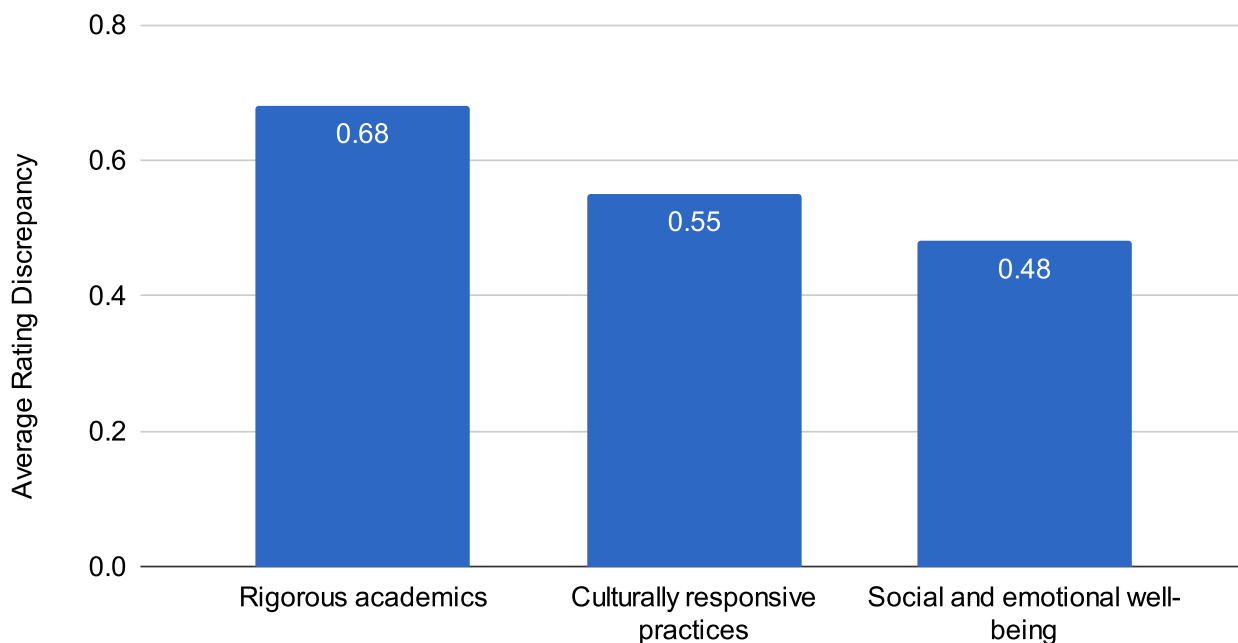
From both rounds of Leading Educators' instructional artifact set analysis in March and May 2021, including ratings of 19 artifact sets from independent rating by external observer pairs. The tool was revised and shortened for the second round. Observers rated each indicator as "2 – Strong evidence", "1 – Some evidence" or "0 – No evidence."

Moving from instructional practice to instructional experience

Most observation rubrics focus on teacher actions and do not consider other critical aspects of instructional quality, such as the teachers' planning or student actions; this incomplete view hampers the ability of support systems to improve instructional quality.[28] Throughout the year, the importance of considering a more holistic view of instruction emerged as a theme. Before the pandemic, we recognized that our observations frequently focused on teacher actions and/or use of materials, but we were not yet integrating what students are experiencing. Rather than focusing on observations of instructional practice alone, we needed to assess the instructional experience, including what students engage with, learn from, and contribute to each lesson.

At the start of the 2020–2021 school year, we asked advisors about how to shift our strategies to be responsive to the moment. Our advisors pushed us to prioritize measures that captured how students were experiencing the new modes of instruction and their virtual learning environments. Similarly, we heard from partners that their priorities were shifting rapidly, and they wanted to know more about student engagement.

After we drafted our first version of our analysis tool aligned to *Teaching for Equity*, we asked advisors to review the preliminary report and tool and make recommendations for improvement. In their feedback, advisors said that what felt most missing from our draft tool was the student experience. For example, we assessed whether the teacher incorporated culturally relevant content and planned for student choice, but we did not identify if students saw relevant current and historic representations of their culture or participated in balanced ways across lines of difference. Advisors urged us to consider how we could reframe the questions we asked and the evidence we focused on to center on students. Without a broader consideration of student answers, participation, and perspectives, how could we understand the instructional experience as a whole?

We focused on sets of artifacts that together could provide a multifaceted glimpse into practice from a variety of evidence sources: annotated lesson plans, videos, and student work. When we revised our analysis tool, we drafted indicators from both the teacher and student perspective, and we also created a student and teacher survey to provide complementary data. One observer reflected, "Starting with the student perspective put me more in the mind-frame of what the students were experiencing than just what the teacher intended, and that definitely changed my responses."

Problem	Potential mitigation strategies
A teacher's intent may differ greatly from the impact on students. Teacher actions alone are too narrow to fully understand the instructional experience, especially when it comes to indicators related to social-emotional well-being or culturally responsive instruction.	• Use instructional artifact sets to more deeply understand what the teacher planned, what the teacher did, and how the students experienced that learning • Consider and strategically select the range of possible evidence sources that could provide a perspective on the indicator, including teacher, student, and an external observer • Prioritize the highest leverage evidence sources to balance efficiency with a depth of understanding

---

28 Cohen, J., Hutt, E., Berlin, R., & Wiseman, E. (2020). The change we cannot see: Instructional quality and classroom observation in the era of common core. *Educational Policy*, 0895904820951114, 8

Narrow and contextualized data collection

While teachers and school partners have a strong interest in better understanding student experiences and learning, our school partners often report feeling overwhelmed by the frequency of data collection activities, including observations, surveys, and standardized tests. Each additional data collection activity creates an additional time commitment for teachers and students. Consequently, each activity must be carefully selected to ensure the utility is worth the time and effort. Education research frequently takes data from communities most impacted by systemic injustices and racism without considering power dynamics around knowledge construction and long-term benefits to those communities.[29] As we recommend something new, we must remain mindful of that human cost and ensure the purpose justifies that cost. The theme of constantly returning to and sharpening the purpose behind data echoed throughout our advisory board meetings.

Early on in the pandemic, advisors recommended we meet our partners where they were by aligning to their priorities and data sources. Instead of additional data collection, they recommended we focus on embedding data gathering efforts within existing programming. While we wanted to understand the impact of our work, students, teachers, and leaders are pulled in many directions during and after the pandemic. We must justify any data collection with a narrower articulation of the rationale and role that evidence will play in supporting change for students.

To shape purpose, advisors recommended bringing in not just district leaders but school-based leaders as co-creators of evaluations of instruction. This could help ensure that data collection would not just serve the professional learning partner but truly drive meaningful changes in classrooms, and that those changes could endure.

Problem	Potential mitigation strategies
Schools and systems are overwhelmed by data collection, and each additional initiative adds a burden to teachers and students. This burden is disproportionately placed on communities most impacted by systemic racism and injustices.	• Co-construct a specific Theory of Action with all partners in the instructional improvement effort • Identify a narrow, primary purpose for data collection in the theory of action • Using the theory of action and purpose, identify a narrow set of indicators and data sources that are most connected to purpose and the theory of action

29 Chicago Beyond (2019). *Why Am I Always Being Researched.* https://chicagobeyond.org/researchequity/

Looking back, looking forward

Our partnership with our Evaluation Advisory Board highlighted the value of leveraging external perspectives and bringing together research and practice experts. Advisor and staff reflections offer additional insights on the key lessons learned through the partnership. Both advisors and staff reflected on the sense of community created through our collaborative work of recognizing and tackling common challenges with diverse perspectives. Staff noted that the ongoing arc of learning created more continuity than one-off collaborations typically afford and provided opportunities to dive deeper in our explorations. Advisors considered that even the most well-researched tools may not yet be closely attuned enough to the relationship between effective teaching practices and equitable teaching practices or between student experience and teacher actions. Academics, professional learning organizations, and the schools and communities we serve all have a role to play in addressing these gaps. As this collaboration demonstrates, more conversation and connections among the three groups can help chart the path forward.

At the same time, advisors urged us to continue narrowing our focus. Using a broad and integrated lens to look at equity in instruction is critical. At the same time, that lens could compromise focus and feasibility of change efforts. Each additional construct or data point adds complexity to both measurement and programming work, whether that be through the additional constructs and evidence sources or the accumulation of the various ideas or frameworks to which we may ask teachers to attend. Without a narrow purpose, any instructional improvement initiative will be pulled in too many directions. We must support partners to find a narrow focus that does not silo content and academics, culturally responsive pedagogies, or social emotional well-being in a way that reinforces inequities and ultimately does harm to all students.

We recognize that equitable classrooms and schools represent just one part of a vast, complex, and fundamentally inequitable educational system. Although this work directly targeted classrooms as the main sites of change, we acknowledge that this is only one of multiple levels at which critical change work must happen. Going forward, we plan to expand our advisory board to incorporate both a focus on measurement and programming decisions. We will also directly incorporate students and community members into the board. Finally, we plan to expand our frame beyond classrooms to better encompass the schools and systems in which they are embedded.

As we look forward, we pose a few important questions for the coming school year:

- How might we examine and address systemic inequities in addition to what happens in classrooms?

- How might we strengthen the specificity of indicators to coherently describe instruction that integrates academics, culturally responsive pedagogies, and social and emotional learning?

- How might we more authentically capture the student experience and what students are learning?

- How might we ensure that instructional improvement and data collection efforts are sufficiently narrow and are useful to stakeholders for supporting change?

- How are our efforts impacting students?

Given the urgency of establishing equitable schools for all students and the new opportunities for systemic change afforded by historic federal investments in public education, collaboration and sharing of practices that make a difference cannot wait.

Leading Educators

f 📷 in 🐦

Leading Educators is reinventing professional development for educators, transforming schools into equity-centered environments where teachers have strong opportunities to learn and practice, in collaboration with others, to bring engaging lessons to life. When we invest in educators, and empower them to lead other educators, we ignite exponential impact across entire school systems, ensuring consistently excellent teaching for every student in every classroom, day after day, year after year.

Leading Educators
Potential, ignited.